Structure
Previews

Polo Boxes Come out of the Crypt: A New View of PLK Function and Evolution

Swadhin Chandra Jana,[1,*] J. Fernando Bazan,[2,*] and Mónica Bettencourt Dias[1,*]
[1]Instituto Gulbenkian de Ciência, Rua da Quinta Grande 6, 2780-156 Oeiras, Portugal
[2]4th & Aspen Life Sciences, 924 4th Street N., Stillwater, MN 55082, USA
*Correspondence: sjana@igc.gulbenkian.pt (S.C.J.), jfbazan@gmail.com (J.F.B.), mdias@igc.gulbenkian.pt (M.B.D.)
http://dx.doi.org/10.1016/j.str.2012.10.008

Polo-like kinases (PLKs) are marked by C-terminal polo box modules with critical protein interaction and subcellular targeting roles. Slevin et al. in this issue of *Structure* reveal the architecture of a hidden set of polo boxes from the divergent PLK4, a critical player in centrosome duplication, shedding new light on the evolution of PLKs and their functionally related kinase ZYG-1.

Polo-like kinases (PLKs), well known for their multiple roles in interphase and mitosis, possess a characteristic modular architecture that joins an N-terminal kinase with a unique C-terminal structure called the polo box domain (PBD), which is comprised of a tandem set of ~100 amino acid polo boxes (PB1 and PB2; Figure 1A). The PBD of PLK1 is involved in substrate binding, regulation of kinase activity, and localization at different subcellular structures and time points. It does so by binding to specific phosphorylated sequences on PLK1 substrates and regulators (reviewed in Archambault and Glover, 2009).

SAK/PLK4 is the most divergent of all PLKs (Carvalho-Santos et al., 2010). This kinase localizes to centrosomes, is required for centrosome duplication, and abnormal protein levels are associated with tumorigenesis (reviewed in Brito et al., 2012). The domain architecture of PLK4 appears distinct from PLK1-3; while PLK1 carries two PBs, PLK4 features only a divergent PB at its C terminus (hereafter called PB3; Figure 1A; Leung et al., 2002). Comparison of the crystal structures of PLK1 PBD and the PLK4 PB3 surprisingly showed that they form different complexes in spite of a conserved modular fold, composed of a twisted antiparallel β sheet packed against a C-terminal helix. Whereas the tandem pair of PLK1 PBs form an intramolecular heterodimer (Figure 1B), the single PLK4 PB3 self-associates as an intermolecular, domain-swapped homodimer (Figure 1D) (Leung et al., 2002; Elia et al., 2003; Cheng et al., 2003). Loss of PLK4 PB3 does not impair centrosome targeting or lead to complete loss of activity, suggesting the presence of other regulatory domains (Leung et al., 2002; Brito et al., 2012; Archambault and Glover, 2009).

Interestingly, a centrally-located, conserved region in PLK4 was shown to dimerize and be sufficient for centrosome localization, and it was also required for PLK4 activity in centrosome formation (Leung et al., 2002; Brito et al., 2012; Slevin et al., 2012). These functional analogies to PBs gave rise to the cryptic polo box (CPB) label (Swallow et al., 2005). Recently, PLK4 was suggested to be regulated by *trans* auto-phosphorylation within the dimer in a degron motif located N-terminal to the CPB domain (Guderian et al., 2010). Moreover, the CPB segment was shown to be essential for Asterless (Asl)/CEP152 binding and ensuing centrosome targeting (reviewed in Brito et al., 2012). Solving the structure of the PLK4 CPB has become crucial for understanding the mechanisms of PLK4's mode of action and deciphering its relationship with the other PLKs.

In this issue of *Structure*, Slevin et al. (2012) present the crystal structure of the *Drosophila* PLK4 CPB at 2.3 Å resolution and reveal a pair of polo boxes, PB1 and PB2, that individually retain the distinctive fold of PLK1 PBs but lack the ~15 residue linker between modules. As a result, the tightly fused CPB polo boxes adopt a head-to-tail packing arrangement (Figure 1C), which dramatically differs from the side-by-side placement of the PLK1 PBs (Figure 1B), or the intertwined homodimer of PLK4 PB3s (Figure 1D). Furthermore, the authors also found that PLK4 CPB forms a stable pseudosymmetric homodimer along its length, clustering their respective N-termini (Slevin et al., 2012). Adding to previous findings in *Drosophila* and human cells (reviewed in Brito et al., 2012), the CPB cassette is required for binding the centriolar protein Asl in vivo and centriole targeting (Slevin et al., 2012). Moreover, the CPB is also necessary to ensure proper PLK4 degradation (Slevin et al., 2012).

The arrangement of PBs in the CPB illustrates the unusual plasticity of binary PB interactions that create modular scaffolds with potentially distinct binding and oligomerization functions. For instance, the two PLK1 PBs and the three PLK4 PBs demonstrate four different organizational modes (Figures 1B–1D and 2A): (1) the PBD of PLK1 (composed of PB1 and PB2), (2) the CPB of PLK4 (PB1 and PB2), (3) the dimer interface of the CPBs of PLK4, and (4) the domain-swapped dimer of PLK4 PB3s (Leung et al., 2002; Slevin et al., 2012). The distinct PB packing modes imposed on the PLK1 PBD and PLK4 CPB by their respective tether lengths has the effect in the former structure of creating a β sheet-protected pocket that binds a Ser/Thr-phosphopeptide chain by a β strand addition mechanism to the N-terminal edge of the PB1 module β sheet. By contrast, the PBs of the CPB create a groove at their modular interface, which is an attractive protein binding site (where the PB1 module of the CPB could also employ a β strand recognition scheme) for interactors, such as Asl, and still undefined substrates (Figures 1C and 2D; Slevin et al., 2012); the absence of a phosphate-recognition cavity in the CPB groove argues that protein binding might be phospho-independent (Slevin et al., 2012).

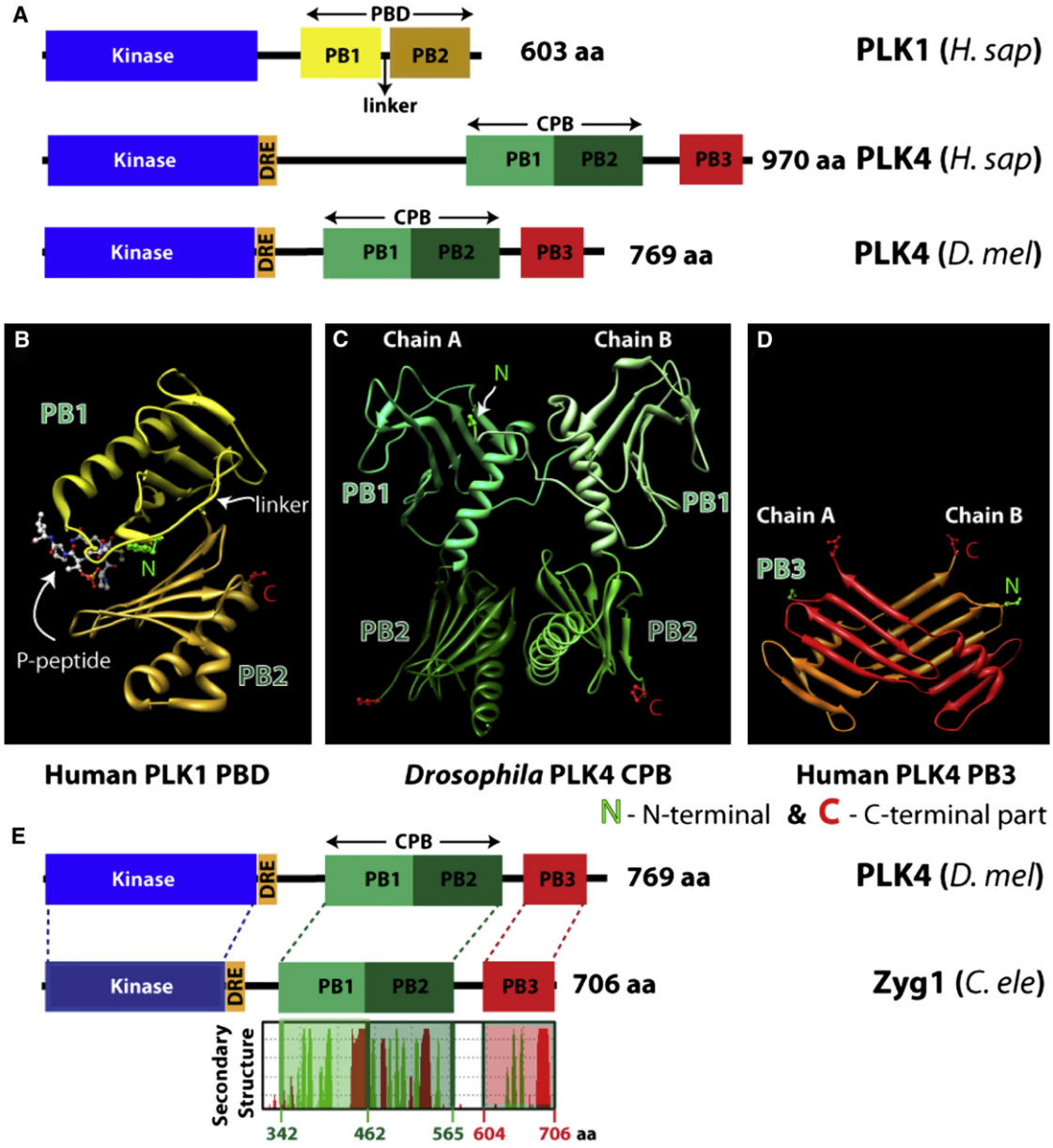

Figure 1. Organizations of PBs in the Different PLKs

(A) Domain architecture of human PLK1 and PLK4 and *Drosophila* PLK4. While PLK1 contains two PBs, PLK4 contains three PBs and a degron regulated by β-TRCP/Slimb (DRE).

(B) Quaternary structures of the PLK1 polo box domain (PBD)/phosphopeptide complex (PDB ID 1UMW).

(C) Dimeric Plk4 cryptic polo box (PDB ID 4G7N).

(D) PLK4 domain-swapped PB3 dimer (PDB ID IMBY).

(E) *Drosophila* PLK4(SAK) and *C. elegans* ZYG-1-predicted kinase (blue), CPB (green), and C-terminal PB3 (red) domains. In ZYG-1, we predict the presence of a CPB and a PB3, which are embedded in the structured C-terminal region (with red and green peaks respectively indicating strongly predicted helices and β strands by PsiPRED, respectively). The repeating structural motif of six β strands and one α helix characteristic of the ~100 PB module is triply noted, though no crystal structure has been generated for ZYG-1.

The PB3 domain plays a dimerization role in PLK4 by contributing to an extensive dimer interface (Figures 1D and 2A), and we speculate that it may also seed the formation of higher-order scaffolds by inter-dimer linking of PB3 modules (Figure 2C). These complexes could, for example, have a structural role in centriole assembly, as recently shown for oligomers of SAS6 that template the 9-fold symmetrical cartwheel structure of centrioles (Figure 2C) (reviewed in Brito et al., 2012).

The addition of CPB to the collection of PB folds expands our view of the structural and functional plasticity of this protein module, which is proving more versatile than other interaction modules tethered to kinase domains (Jin and Pawson, 2012). The CPB notably stabilizes the PLK4 dimer and juxtaposes the kinase domains, which facilitates the

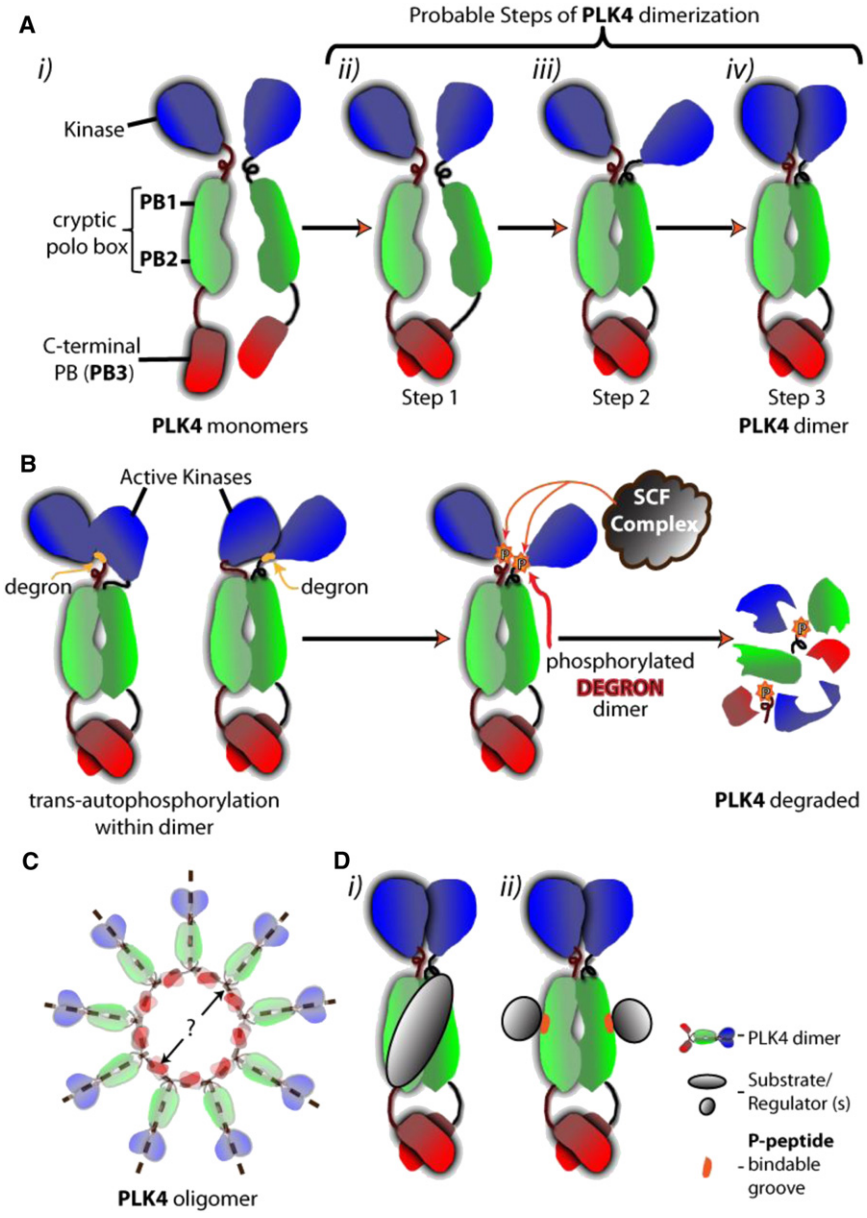

Figure 2. Speculative Model for Quaternary Structures of PLK4 and Their Modes of Action

(A) Schematic presentation of possible steps of PLK4 dimerization (see main text for references). We speculate that PLK4 may attain different states of dimerization through different intermediate steps, an example of which is provided here (ii–iv). Two monomers (i) may dimerize through C-terminal intertwining of PB3s (ii); pseudosymmetric dimerization of the CPB can also lead to PLK4 dimer formation (iii). Lastly, PLK4 kinase domains pack together (PDB ID 3COK) to dimerize (iv), unlike the case of PLK1.

(B) Trans-autophosphorylation of the "degron" motif of one PLK4 subunit by the active kinase domain of the other molecule within the dimer is likely to precede PLK4 degradation. Subsequently, the SCF complex recognizes its substrate motif, the phosphorylated "degron" dimer, leading to PLK4 degradation.

(C) We speculate that, instead of the intermolecular embrace, the PB3 domains of a PLK4 dimer could oligomerize with neighboring PB3s of other dimers and associate into ring-like assemblies of different topologies. These oligomers could, for example, resemble a cartwheel, placing the N-terminal kinase domains radially outward.

(D) The packed PB1-PB2 domains form large, concave groove(s) that could accommodate the chains of PLK4 interactors (i–ii), such as Asl, and yet undiscovered substrates/regulators.

trans autophosphorylation of both degrons. Simultaneously, the CPB creates a binding platform for globular substrates like Asl (Figures 2B and 2D). The precise mutational ablation of the different dimer interfaces in PLK4 will drive a better understanding of how PLK4 function is regulated as both a catalytic entity and architectural scaffold.

PLK4s are the outliers of the PLK family, as judged by their greater divergence of both kinase domains and PB arrays (Carvalho-Santos et al., 2010). As Slevin et al.'s work now demonstrates, the architecture of the CPB (coupled to the oligomerizing PB3 domain) provides the PLK4 kinase with a unique scaffolding ability absent in the other PLKs and sets PLK4 functionally apart as an enzymatic and perhaps also structural component of the centriole (Figure 2).

A PLK4 ortholog has not been detected in the *C. elegans* genome (Carvalho-Santos et al., 2010). There is an intriguing molecule, ZYG-1, that operationally fits the mold of an active PLK4, including how it is degraded, but its kinase domain lies on a distal branch of the kinome tree from PLKs and is thought to be a case of functional convergence (Carvalho-Santos et al., 2010, Peel et al., 2012). However, the featureless 435 residues following the ZYG-1 kinase domain are rich in predicted secondary structure, and, using sensitive fold recognition methods (Bazan and de Sauvage, 2009) focused on this C-terminal globular region of ZYG-1, we can predict the presence of a PLK4-like array of CPB and PB3 domains (Figure 1E). If the CPB and PB3 are vital to this PLK4-like outlier, then perhaps ZYG-1 is an unorthodox PLK4 whose kinase sequence has evolutionarily drifted more rapidly than the main PLK4 branch, obscuring its true orthology. PLK4 phylogeny is presently limited to fungi and animals (Carvalho-Santos et al., 2010); it will be important to investigate the divergence of CPB and PBD domains in order to further understand PLK4's evolutionary origins and its function in centriole biogenesis.

ACKNOWLEDGMENTS

We apologize to colleagues whose work was not discussed or cited owing to space constraints. We thank Inês Cunha-Ferreira, Zita Carvalho-Santos, Tiago M. Bandeiras, Gregory C. Rogers, Jose Pereira-Leal, and Sihem Zitouni for critically reading the manuscript. S.C.J. and M.B.D. are funded by an EMBO Installation grant (cofunded by FCT and Instituto Gulbenkian de Ciência) and the EMBO YIP Program, (FP7/2010)/ERC Grant "261344-CentrioleStructNumber," and Fundação para a Ciência e Tecnologia (FCT) "PTDC/BIA-BCM/105602/2008".

REFERENCES

Archambault, V., and Glover, D.M. (2009). Nat. Rev. Mol. Cell Biol. *10*, 265–275.

Bazan, J.F., and de Sauvage, F.J. (2009). Cell *138*, 1055–1056.

Brito, D.A., Gouveia, S.M., and Bettencourt-Dias, M. (2012). Curr. Opin. Cell Biol. *24*, 4–13.

Carvalho-Santos, Z., Machado, P., Branco, P., Tavares-Cadete, F., Rodrigues-Martins, A., Pereira-Leal, J.B., and Bettencourt-Dias, M. (2010). J. Cell Sci. *123*, 1414–1426.

Cheng, K.Y., Lowe, E.D., Sinclair, J., Nigg, E.A., and Johnson, L.N. (2003). EMBO J. *22*, 5757–5768.

Elia, A.E., Rellos, P., Haire, L.F., Chao, J.W., Ivins, F.J., Hoepker, K., Mohammad, D., Cantley, L.C., Smerdon, S.J., and Yaffe, M.B. (2003). Cell *115*, 83–95.

Guderian, G., Westendorf, J., Uldschmid, A., and Nigg, E.A. (2010). J. Cell Sci. *123*, 2163–2169.

Jin, J., and Pawson, T. (2012). Philos. Trans. R. Soc. Lond. B Biol. Sci. *367*, 2540–2555.

Leung, G.C., Hudson, J.W., Kozarova, A., Davidson, A., Dennis, J.W., and Sicheri, F. (2002). Nat. Struct. Biol. *9*, 719–724.

Peel, N., Dougherty, M., Goeres, J., Liu, Y., and O'Connell, K.F. (2012). J. Cell Sci. *125*, 3535–3544.

Slevin, L.K., Nye, J., Pinkerton, D.C., Buster, D.W., Rogers, G.C., and Slep, K.C. (2012). Structure *20*, this issue, 1905–1917.

Swallow, C.J., Ko, M.A., Siddiqui, N.U., Hudson, J.W., and Dennis, J.W. (2005). Oncogene *24*, 306–312.

Retroviral Intasomes: Progress and Questions

Min Li[1] and Robert Craigie[1,*]
[1]Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA
*Correspondence: bobc@helix.nih.gov
http://dx.doi.org/10.1016/j.str.2012.10.004

In this issue of *Structure*, Gupta and colleagues apply a combination of biophysical approaches to study the solution properties of prototype foamy virus (PFV) integrase alone and in complex with viral DNA ends (intasome). The results complement and extend previous structural studies of PFV intasomes by X-ray crystallography and highlight the synergy of solution and crystallographic approaches to the study of nucleoprotein complexes.

Integration of retroviral DNA into cellular DNA is an essential step in the replication cycle of retroviruses. The virally encoded integrase protein is the key enzyme that carries out the DNA cutting and joining steps. Viral DNA is synthesized by reverse transcription in the cytoplasm and remains associated with viral and cellular proteins as a large nucleoprotein complex called the preintegration complex (PIC). After transport from the cytoplasm to nucleus, integrase within the PIC catalyzes insertion of the viral DNA into chromosomal DNA. Functional studies of PICs show that integrase is tightly bound to the viral DNA ends, but the low abundance of PICs in extracts of infected cells prevents direct analysis of the structure and organization of the complex. Purified retroviral integrases catalyze integration in vitro in the presence of a divalent metal ion. These reaction systems have proved to be invaluable for studying the biochemical mechanism. Integration occurs by a one step direct transesterification reaction in which the 3′ ends of the viral DNA attack a pair of phosphodiester bonds at the sites of integration into the target DNA (Engelman et al., 1991). The first stable nucleoprotein complex on the integration reaction pathway is the stable synaptic complex (SSC), comprising a pair of viral DNA ends synapsed by a tetramer of integrase (Li et al., 2006). The integration product remains tightly associated with the integrase tetramer in another stable complex called the strand transfer complex (STC). Collectively, these complexes are referred to as intasomes.

Structural studies of retroviral integrases have progressed at a painfully slow pace. The structures of the individual domains of several integrases were determined by the mid 1990s, and two domain structures were later determined (Chiu and Davies, 2004). Apart from the conserved catalytic core domain dimer interface, the relative positions of the domains differed among the structures, highlighting the flexibility of the domain linkers in the absence of DNA. The major obstacle to structural studies is the propensity of integrases to aggregate, especially in the presence of viral DNA substrate and under conditions where the enzyme is active. The first intasome structures were not determined until 2010 when Cherepanov and colleagues solved the structure of prototype foamy virus (PFV) SSC and STC intasomes (Hare et al., 2010; Maertens et al., 2010). This major advance was facilitated by the vastly superior properties of the PFV enzyme compared with the more extensively studies HIV-1 and avian counterparts. Unlike the HIV-1 and avian integrases, PFV integrase efficiently forms homogeneous SSC on short oligonucleotides with the sequence of the viral DNA ends and is highly active in vitro. The PFV intasomes are currently the only retroviral intasomes for which crystal structures have been determined and are therefore the best template for modeling HIV-1 intasomes. High-resolution structures of the HIV-1 intasomes are required because current inhibitors of integrase

